

## Figure 1A

Identity to SeqID No:1 is indicated by a dot, and a dash (“-“) indicates a missing nucleotide.

	*	20	*	40	*	
Seqid1 :	ATGAGATCTTCTTTCCGGTTGAAGCCGATTGTGTTTTATCTTATGGGTGT					: 50
Seqid3 :	.....					: 50
Seqid5 :	.....C.....					: 50

	60	*	80	*	100	
Seqid1 :	TATGCTATATCATCATAGTTATGCCGAAGATGCAGGGCGCGCGGGCAGCG					: 100
Seqid3 :	.....					: 100
Seqid5 :	..C.....T.....					: 100

	*	120	*	140	*	
Seqid1 :	AGGCGCAGATACAGGTTTTGGAAGATGTGCACGTCAAGGCGAAGCGCGTA					: 150
Seqid3 :	.....					: 150
Seqid5 :	.....					: 150

	160	*	180	*	200	
Seqid1 :	CCGAAAGACAAAAAGTGTTTACCGATGCGCGTGCCGTATCGACCCGTCA					: 200
Seqid3 :	.....					: 200
Seqid5 :	.....					: 200

	*	220	*	240	*	
Seqid1 :	GGATATATTCAAATCCAGCGAAAACCTCGACAACATCGTACGCAGCATCC					: 250
Seqid3 :	.....					: 250
Seqid5 :	.....					: 250

	260	*	280	*	300	
Seqid1 :	CCGGTGCGTTTACACAGCAAGATAAAAGCTCGGGCATTGTGTCTTTGAAT					: 300
Seqid3 :	.....					: 300
Seqid5 :	.....					: 300

**Figure 1B**

	*	320	*	340	*	
Seqid1	:	ATTCGCGGCGACAGCGGGTTCGGGCGGGTCAATACGATGGTGGACGGCAT	:	350		
Seqid3	:	.....	:	350		
Seqid5	:	.....	:	350		
		360	*	380	*	400
Seqid1	:	CACGCAGACCTTTTATTCGACTTCTACCGATGCGGGCAGGGCAGGCGGTT	:	400		
Seqid3	:	.....	:	400		
Seqid5	:	.....	:	400		
		*	420	*	440	*
Seqid1	:	CATCTCAATTCGGTGCATCTGTGCGACAGCAATTTTATTGCCGGACTGGAT	:	450		
Seqid3	:	.....	:	450		
Seqid5	:	.....	:	450		
		460	*	480	*	500
Seqid1	:	GTCGTCAAAGGCAGCTTCAGCGGCTCGGCAGGCATCAACAGCCTTGCCGG	:	500		
Seqid3	:	.....	:	500		
Seqid5	:	.....	:	500		
		*	520	*	540	*
Seqid1	:	TTCGGCGAATCTGCGGACTTTAGGCGTGGATGACGTCGTTTCAGGGCAATA	:	550		
Seqid3	:	.....	:	550		
Seqid5	:	.....	:	550		
		560	*	580	*	600
Seqid1	:	ATACCTACGGCCTGCTGCTAAAAGGTCTGACCGGCACCAATTCAACCAAA	:	600		
Seqid3	:	.....	:	600		
Seqid5	:	.....	:	600		
		*	620	*	640	*
Seqid1	:	GGTAATGCGATGGCGGCGATAGGTGCGCGCAAATGGCTGGAAAGCGGAGC	:	650		

Figure 1C

```
Seqid3 : ..... : 650
Seqid5 : ..... : 650

        660          *          680          *          700
Seqid1 : ATCTGTCGGTGTGCTTTACGGGCACAGCAGGCGCACGTGGGCGCAAAATT : 700
Seqid3 : .....GCGT..... : 700
Seqid5 : .....GCGT..... : 700

        *          720          *          740          *
Seqid1 : ACCGCGTGGGCGGGCGGGCAGCACATCGGAAATTTTGGCGCGGAATAT : 750
Seqid3 : ..... : 750
Seqid5 : ..... : 750

        760          *          780          *          800
Seqid1 : CTGGAACGGCGCAAACAGCGATATTTTGTACAAGAAGCGGGTTGAAATT : 800
Seqid3 : .....G..... : 800
Seqid5 : T.....G.....G..T.CT..... : 800

        *          820          *          840          *
Seqid1 : CAATTCCAACAGCGGAAATGGGAGCGGGATTTCCAAAGGCCGTACTGGA : 850
Seqid3 : ..... : 850
Seqid5 : .....G.....A.....AAC.G... : 850

        860          *          880          *          900
Seqid1 : AAACCAAGTGGTATCAAAAATACAATGACCCCCAAGAACTGCAAAAATAC : 900
Seqid3 : ..... : 900
Seqid5 : ..TA...CC...A...T...CA---.....A..... : 897

        *          920          *          940          *
Seqid1 : ATCGAAGGTCATGACAAAAGCTGGCGGGAAAACCTGGCGCCGCAATACGA : 950
Seqid3 : ..... : 950
Seqid5 : ..... : 947
```

# Figure 1D

	960	*	980	*	1000	
Seqid1	:	CATCACCCCATCGATCCGTCCAGCCTGAAGCAGCAGTCGGCAGGCAATC	:	1000		
Seqid3	:	.....	:	1000		
Seqid5	:	.....	:	997		

	*	1020	*	1040	*	
Seqid1	:	TGTTTAAATTGGAATACGACGGCGTATTCAATAAATACACGGCGCAATTT	:	1050		
Seqid3	:	.....	:	1050		
Seqid5	:	.....	:	1047		

	1060	*	1080	*	1100	
Seqid1	:	CGCGATTTTAAACACCAAAATCGGCAGCCGCAAAATCATCAACCGCAATTA	:	1100		
Seqid3	:	.....	:	1100		
Seqid5	:	.....	:	1097		

	*	1120	*	1140	*	
Seqid1	:	TCAATTCAATTACGGTTTATCTTTAAACTCATATGCCAACCTCAATCTGA	:	1150		
Seqid3	:	.....	:	1150		
Seqid5	:	...G.....G....G...C.G...A.....	:	1147		

	1160	*	1180	*	1200	
Seqid1	:	CCGCAGCCTACAATTCGGGCAGGCAGAAATATCCGAAAGGGTCGAAGTTT	:	1200		
Seqid3	:	.....	:	1200		
Seqid5	:	.....	:	1197		

	*	1220	*	1240	*	
Seqid1	:	ACAGGCTGGGGGCTTTTAAAAGATTTTGAAACCTACAACAACGCGAAAAT	:	1250		
Seqid3	:	.....	:	1250		
Seqid5	:	.....G.....	:	1247		

	1260	*	1280	*	1300	
Seqid1	:	CCTCGACCTCAACAACACCGCCACCTTCCGGCTGCCCCGCGAAACCGAGT	:	1300		
Seqid3	:	.....	:	1300		
Seqid5	:	.....	:	1297		

## Figure 1E

```

          *      1320          *      1340          *
Seqid1 : TGCAAACCAC TTTGGGCTTCAATTATTTCCACAACGAATACGGCAAAAAC : 1350
Seqid3 : ..... : 1350
Seqid5 : ..... : 1347

```

```

          1360          *      1380          *      1400
Seqid1 : CGCTTTCCTGAAGAATTGGGGCTGTTTTTCGACGGTCCGGATCAGGACAA : 1400
Seqid3 : ..... : 1400
Seqid5 : .....T..... : 1397

```

```

          *      1420          *      1440          *
Seqid1 : CGGGCTTTATTCTTATTTGGGGCGGTTTAAGGGCGATAAAGGGCTGCTGC : 1450
Seqid3 : ..... : 1450
Seqid5 : ..... : 1447

```

```

          1460          *      1480          *      1500
Seqid1 : CCCAAAAATCAACCATCGTCCAACCGGCCGCGCAGCCAATATTTCAACACG : 1500
Seqid3 : ..... : 1500
Seqid5 : .....T..... : 1497

```

```

          *      1520          *      1540          *
Seqid1 : TTCTACTTCGATGCCGCGCTCAAAAAAGACATTTACCGCTTAAACTACAG : 1550
Seqid3 : ..... : 1550
Seqid5 : ..... : 1547

```

```

          1560          *      1580          *      1600
Seqid1 : CACCAATACCGTCGGCTACCGTTTCGGCGGCGAATATACGGGCTATTACG : 1600
Seqid3 : ..... : 1600
Seqid5 : ..... : 1597

```

```

          *      1620          *      1640          *
Seqid1 : GCTCGGATGACGAATTTAAGCGGGCATTCGGAGAAAACTCGCCGACATAC : 1650

```

**Figure 1F**

Seqid3 : ..... : 1650  
Seqid5 : ..... : 1647

1660 \* 1680 \* 1700  
Seqid1 : AAGAAACATTGCAACCAGAGCTGCGGAATTTATGAACCCGTATTGAAAAA : 1700  
Seqid3 : ..... : 1700  
Seqid5 : .....G.....G..... : 1697

\* 1720 \* 1740 \*  
Seqid1 : ATACGGCAAAAAGCGCGCCAACAACCATTCGGTCAGCATTAGTGCGGACT : 1750  
Seqid3 : ..... : 1750  
Seqid5 : ..... : 1747

1760 \* 1780 \* 1800  
Seqid1 : TCGGCGATTATTTTCATGCCGTTGCCAGCTATTCGCGCACACACCGTATG : 1800  
Seqid3 : ..... : 1800  
Seqid5 : ..... : 1797

\* 1820 \* 1840 \*  
Seqid1 : CCCAACATCCAAGAAATGTATTTTTCCCAAATCGGCGACTCCGGCGTTCA : 1850  
Seqid3 : ..... : 1850  
Seqid5 : ..... : 1847

1860 \* 1880 \* 1900  
Seqid1 : CACCGCCTTAAAACCAGAGCGCGCAAACACTTGGCAATTTGGCTTCAATA : 1900  
Seqid3 : ..... : 1900  
Seqid5 : ..... : 1897

\* 1920 \* 1940 \*  
Seqid1 : CCTATAAAAAAGGATTGTTAAAACAAGATGATACATTAGGATTAAACTG : 1950  
Seqid3 : ..... : 1950  
Seqid5 : ..... : 1947

# Figure 1G

	1960	*	1980	*	2000	
Seqid1	:	GTCGGCTACCGCAGCCGCATCGACA	ACTACATCCACAACGTTTACGGGAA	:	2000	
Seqid3	:	.....		:	2000	
Seqid5	:	.....		:	1997	

	*	2020	*	2040	*	
Seqid1	:	ATGGTGGGATTTGAACGGGAATATTCCGAGCTGGGTCAGCAGCACCGGGC	:	2050		
Seqid3	:	.....	:	2050		
Seqid5	:	.....G.....	:	2047		

	2060	*	2080	*	2100	
Seqid1	:	TTGCCTACACCATCCAACACCGCAATTTCAAAGACAAAGTACACAAACAC	:	2100		
Seqid3	:	.....	:	2100		
Seqid5	:	.....T.....G.....	:	2097		

	*	2120	*	2140	*	
Seqid1	:	GGTTTTGAGTTGGAGCTGAATTACGATTATGGGCGTTTTTTCACCAACCT	:	2150		
Seqid3	:	.....	:	2150		
Seqid5	:	.....	:	2147		

	2160	*	2180	*	2200	
Seqid1	:	TTCTTACGCCTATCAAAAAAGCAGCAACCGACCAACTTCAGCGATGCGA	:	2200		
Seqid3	:	.....	:	2200		
Seqid5	:	.....	:	2197		

	*	2220	*	2240	*	
Seqid1	:	GCGAATCGCCCCAACAATGCGTCCAAAGAAGACCAACTCAAACAAGGTTAT	:	2250		
Seqid3	:	.....	:	2250		
Seqid5	:	.....	:	2247		

	2260	*	2280	*	2300	
Seqid1	:	GGGTTGAGCAGGGTTTCCGCCCTGCCGCGAGATTACGGACGTTTGGAAGT	:	2300		
Seqid3	:	.....	:	2300		

## Figure 1H

Seqid5 : ..... : 2297

                  \*          2320                  \*          2340                  \*

Seqid1 : CGGTACGCGCTGGTTGGGCAACAACTGACTTTGGGCGGCGCGATGCGCT : 2350

Seqid3 : ..... : 2350

Seqid5 : ..... : 2347

                  2360                  \*          2380                  \*          2400

Seqid1 : ATTTTCGGCAAGAGCATCCGCGCGACGGCTGAAGAACGCTATATCGACGGC : 2400

Seqid3 : ..... : 2400

Seqid5 : ..... : 2397

                  \*          2420                  \*          2440                  \*

Seqid1 : ACCAACGGGGGAAATACCAGCAATGTCCGGCAACTGGGCAAGCGTTCCAT : 2450

Seqid3 : ..... : 2450

Seqid5 : .....T..... : 2447

                  2460                  \*          2480                  \*          2500

Seqid1 : CAAACAAACCGAAACCCTTGCCCGCCAGCCTTTGATTTTTGATTTTTACG : 2500

Seqid3 : ..... : 2500

Seqid5 : .....T..... : 2497

                  \*          2520                  \*          2540                  \*

Seqid1 : CCGCTTACGAGCCGAAGAAAAACCTTATTTTCCGCGCCGAAGTCAAAAAT : 2550

Seqid3 : ..... : 2550

Seqid5 : ..... : 2547

                  2560                  \*          2580                  \*          2600

Seqid1 : CTGTTTCGACAGGCGTTATATCGATCCGCTCGATGCGGGCAATGATGCGGC : 2600

Seqid3 : ..... : 2600

Seqid5 : ..... : 2597



## Figure 1I

                  \*          2620                  \*          2640                  \*

Seqid1 : AACGCAGCGTTATTACAGTTCGTTTCGACCCGAAAGACAAGGACGAAGAAG : 2650

Seqid3 : ..... : 2650

Seqid5 : .....C.....C. : 2647

                  2660                  \*          2680                  \*          2700

Seqid1 : TAACGTGTAATGCTGATAAAACGTTGTGCAACGGCAAATACGGCGGCACA : 2700

Seqid3 : ..... : 2700

Seqid5 : ..... : 2697

                  \*          2720                  \*          2740                  \*

Seqid1 : AGCAAAAGCGTATTGACCAATTTTGCACGCGGACGCACCTTTTGTGATAAC : 2750

Seqid3 : ..... : 2750

Seqid5 : ..... : 2747

                  2760

Seqid1 : GATGAGCTACAAGTTTTAA : 2769

Seqid3 : ..... : 2769

Seqid5 : ..... : 2766

## Figure 2A

Identity to SeqID No:2 is indicated by a dot, and a dash ("-") indicates a missing amino acid.

```

          *          20          *          40          *
Seqid2 : MRSSFRLKPICFYLMGVMLYHHSYAEDAGRAGSEAQIQVLEDVHVKA KRV : 50
Seqid4 : ..... : 50
Seqid6 : .....T...Y..... : 50

```

```

          60          *          80          *          100
Seqid2 : PKDKKVFTDARAVSTRQDIFKSSSENLDNIVRSIPGAFTQQDKSSGIVSLN : 100
Seqid4 : ..... : 100
Seqid6 : ..... : 100

```

```

          *          120          *          140          *
Seqid2 : IRGDSGFGRVNTMVDGITQTFYSTSTDAGRAGGSSQFGASVDSNFIAGLD : 150
Seqid4 : ..... : 150
Seqid6 : ..... : 150

```

```

          160          *          180          *          200
Seqid2 : VVKGSFSGSAGINSLAGSANLRTLGVDDVVQGNNTYGLLLKGLTGTNSTK : 200
Seqid4 : ..... : 200
Seqid6 : ..... : 200

```

```

          *          220          *          240          *
Seqid2 : GNAMAAIGARKWLESGASVGVLYGHSRRTWAQNYRVGGGGQHIGNFGAEY : 250
Seqid4 : .....SV..... : 250
Seqid6 : .....SV..... : 250

```

## Figure 2B

```

                260          *          280          *          300
Seqid2 : LERRKQRYFVQEGGLKFNSNSGKWERDFQRPYWKTKWYQKYNDPQELQKY : 300
Seqid4 : ..... : 300
Seqid6 : .....A....D....L..QQ..Y.P.KN..N-..... : 299

```

```

                *          320          *          340          *
Seqid2 : IEGHDKSWREN LAPQYDITPIDPSSLKQQSAGNLFKLEYDGVFNKYTAQF : 350
Seqid4 : ..... : 350
Seqid6 : ..... : 349

```

```

                360          *          380          *          400
Seqid2 : RDLN TKIGSRKI INRNYQFN YGLSLNSYANLNLTAAYNSGRQKYPKGSKF : 400
Seqid4 : ..... : 400
Seqid6 : .....P.T..... : 399

```

```

                *          420          *          440          *
Seqid2 : TGWGLLKDFETYNNAKILDNLNTATFRLPRETELQTTLGFNYFHNEYGKN : 450
Seqid4 : ..... : 450
Seqid6 : ..... : 449

```

```

                460          *          480          *          500
Seqid2 : RFPEELGLFFDGPDQDNGLYSYLGRFKGDKGLLPQKSTIVQPAGSQYFNT : 500
Seqid4 : ..... : 500
Seqid6 : ..... : 499

```

```

                *          520          *          540          *
Seqid2 : FYFDAALKKDIYRLNYSTNTVGYRFGGEYTGYYGSDDEFKRAFGENSPTY : 550
Seqid4 : ..... : 550
Seqid6 : ..... : 549

```

Figure 2C

560                   \*                   580                   \*                   600  
Seqid2 : KKHCNQSCGIYEPVLKKYGKKRANNHVSISADFGDYFMPFASYSRTHRM : 600  
Seqid4 : ..... : 600  
Seqid6 : ....R..... : 599

                 \*                   620                   \*                   640                   \*  
Seqid2 : PNIQEMYFSQIGDSGVHTALKPERANTWQFGFNTYKKGLLKQDDTLGLKL : 650  
Seqid4 : ..... : 650  
Seqid6 : ..... : 649

                 660                   \*                   680                   \*                   700  
Seqid2 : VGYRSRIDNYIHNVYGKWWDLNGNIPSWVSSTGLAYTIQHRNFKDKVHKH : 700  
Seqid4 : ..... : 700  
Seqid6 : .....D..... : 699

                 \*                   720                   \*                   740                   \*  
Seqid2 : GFELELNVDYGRFFTNLSYAYQKSTQPTNFSDASESPNNASKEDQLKQGY : 750  
Seqid4 : ..... : 750  
Seqid6 : ..... : 749

                 760                   \*                   780                   \*                   800  
Seqid2 : GLSRVSALPRDYGRLEVGTWLGKNKLTGGAMRYFGKSIRATAEERYIDG : 800  
Seqid4 : ..... : 800  
Seqid6 : ..... : 799

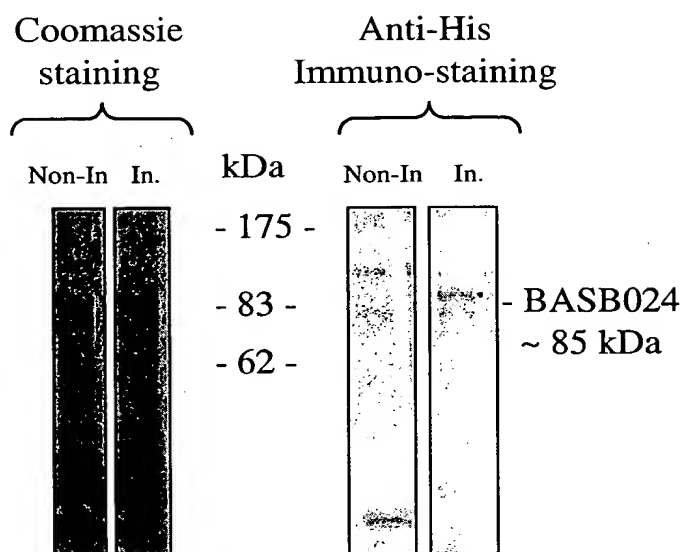
                 \*                   820                   \*                   840                   \*  
Seqid2 : TNGGNTSNVRQLGKRSIKQTETLARQPLIFDFYAAAYEPKKNLIFRAEVKN : 850  
Seqid4 : ..... : 850  
Seqid6 : .....F..... : 849

## Figure 2D

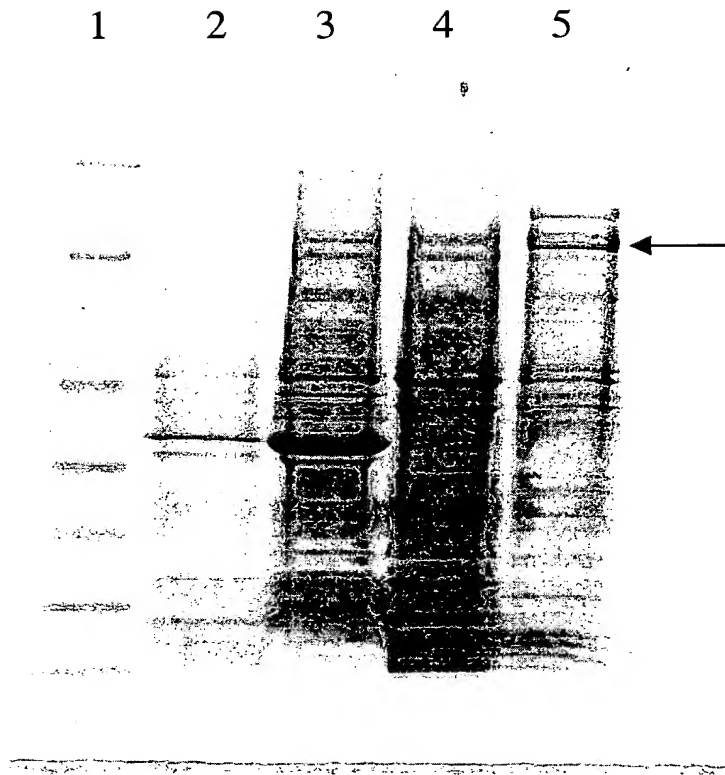
860                   \*                   880                   \*                   900  
Seqid2 : LFD RRYIDPLDAGNDAATQ RYYSSFD PKDKDEEVTCNADKTL CNGKYGGT : 900  
Seqid4 : ..... : 900  
Seqid6 : .....D..... : 899

                  \*                   920  
Seqid2 : SKSVLTNFARGRTFLITMSYKF : 922  
Seqid4 : ..... : 922  
Seqid6 : ..... : 921

**Figure 3**

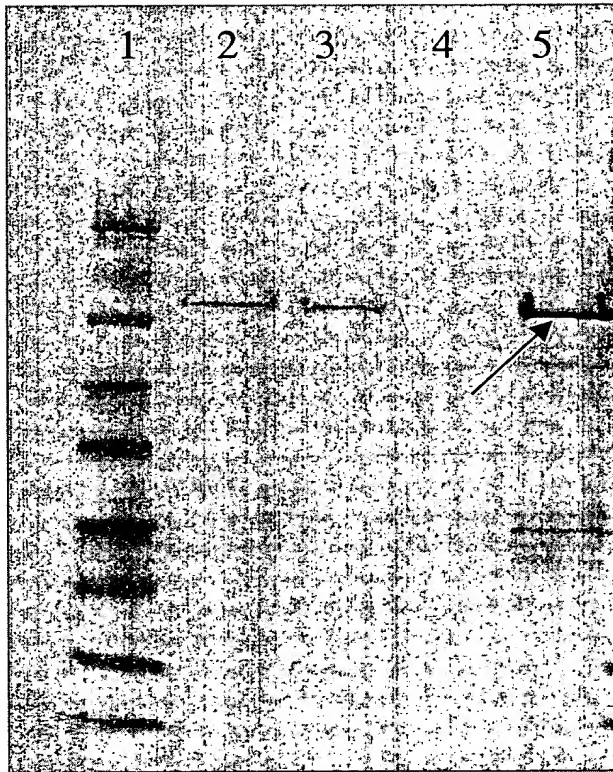


**Figure 4**



Lanes :      1 : MW : 175, 83, 62, 47.5, 32.5, 25, 16.5, 6.5  
              2 : Start  
              3 : Flowthrough  
              4 : 5 mM imidazole pool  
              5 : BASB024 enriched pool

**Figure 5**



Lanes :      1 : MW : 175, 83, 62, 47.5, 32.5, 25, 16.5, 6.5  
              2 : Start  
              3 : Flowthrough  
              4 : 5 mM imidazole pool  
              5 : BASB024 enriched pool



**Figure 6**

